MOLECULAR AND CELLULAR BIOLOGY, Sept. 1985, p. 2487–2490
0270-7306/85/092487-04$02.00/0

Vol. 9, No. 5

NOTES

The Spliced Leader Sequence of *Trypanosoma brucei* Has a Potential Role as a Cap Donor Structure

MICHAEL J. LENARDO,[1]† DAVID M. DORFMAN,[1,2] AND JOHN E. DONELSON[1,2]*

Department of Biochemistry[1] and Ph.D. Program in Genetics,[2] University of Iowa, Iowa City, Iowa 52242

Received 28 February 1985/Accepted 14 June 1985

***Trypanosoma brucei brucei* and other trypanosomatid species are unique among eucaryotes because transcription of their protein-coding genes is discontinuous. The 5′ ends of their mRNAs consist of an identical 35-nucleotide spliced leader which is encoded at a separate locus from that for the body of the protein-coding transcript. We show here that the spliced leader transcript contains a 5′ cap structure and suggest that at least one function of the spliced leader sequence is to provide a cap structure to trypanosome mRNAs.**

Recent studies have shown that transcription of protein-coding genes in *Trypanosoma brucei* is discontinuous (4, 11, 16). The 5′ ends of many, if not all, trypanosome mRNAs contain a conserved 35-nucleotide sequence which is encoded separately in the genome in a 1.35-kilobase gene repeat found in about 200 tandemly arranged copies (6, 7, 19). A transcript of 135 to 141 nucleotides found in steady-state RNA may be a precursor to the 35 nucleotides present in mRNAs (4, 11, 16). The function of the spliced leader sequence, however, remains obscure.

We investigated the structure of the 5′ end of the spliced leader transcript with an enzyme complex (guanylyl transferase) from vaccinia virions; this complex can synthesize cap structures on RNAs that have 5′ di- or triphosphate termini (17). The following reactions (excluding a methylation reaction) are catalyzed:

$$\text{pppN(pN)}_n \rightarrow \text{ppN(pN)}_n + \text{P}_i$$
$$[\alpha\text{-}^{32}\text{P}]\text{GTP} + \text{ppN(pN)}_n \rightarrow \text{G(5')}^{32}\text{pppN(pN)}_n + \text{PP}_i$$

RNAs which bear either a preexisting cap structure or a monophosphate at the 5′ terminus are not substrates. A preexisting cap structure can be removed by periodate oxidation and beta elimination by a weak base (9, 17), resulting in a triphosphate terminus which can be capped by guanylyl transferase.

Total trypanosome RNA was isolated from *T. brucei* IaTat 1.2 (12) and treated with guanylyl transferase (Bethesda Research Laboratories, Inc., Gaithersburg, Md.) and [α-^{32}P]GTP under conditions recommended by the manufacturer. The labeled RNA was hybridized to 3 μg of various plasmids immobilized on nitrocellulose (11). No labeled RNA bound to the genes for alpha- or beta-tubulin, the variable surface glycoprotein (VSG) expressed in these organisms, or the 5.8S rRNA (Fig. 1, Pre). A variable and small amount of labeled RNA bound to the gene for the spliced leader repeat (discussed below). As a control, the 5S rRNA was strongly labeled, consistent with the fact that it has a 5′ polyphosphate terminus (12). Therefore, the 5′ ends of the majority of transcripts for the spliced leader, VSG, and tubulins consist of either a monophosphate or a cap structure.

We then used periodate oxidation and beta elimination to remove cap structures from 24 μg of total trypanosome RNA

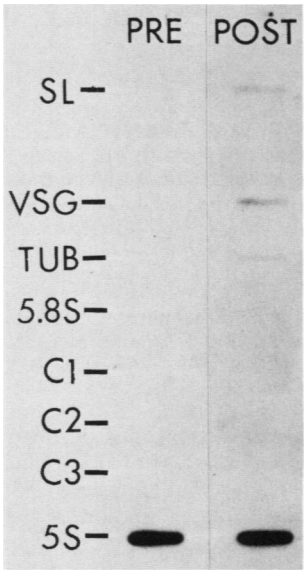

FIG. 1. Autoradiogram of the hybridization to specific genes of trypanosome RNA capped in vitro with [α-^{32}P]GTP. Total trypanosome RNA was capped before (Pre) and after (Post) removal of the preexisting cap structures. Slots contained immobilized trypanosome DNA possessing the following genes or plasmids: cloned spliced leader gene repeat (SL) (7); cloned IaTat 1.2 VSG gene (VSG) (18); cloned tubulin gene repeat including alpha- and beta-tubulin genes (TUB) (10; a gift of Lyn Yarbrough, University of Kansas); cloned 5.8S rRNA gene (5.8S) (8); pUC9 plasmid (C1); pBR322 plasmid (C2); cloned MVAT 4 VSG gene (C3) (13); and cloned 5S gene repeat (5S) (12). Preexisting cap structures were removed from 24 μg of total RNA by chemical degradation according to the procedure of Fraenkel-Conrat and Steinschneider (9). Approximately 2×10^6 to 5×10^6 cpm of labeled RNA from this reaction was hybridized with 3 μg of various genes immobilized on nitrocellulose with a slot manifold (Schleicher & Schuell, Inc., Keene, N.H.) under conditions described previously (11, 20) and autoradiographed as shown.

* Corresponding author.
† Present address: Whitehead Institute for Biomedical Research, Cambridge, MA 02142.

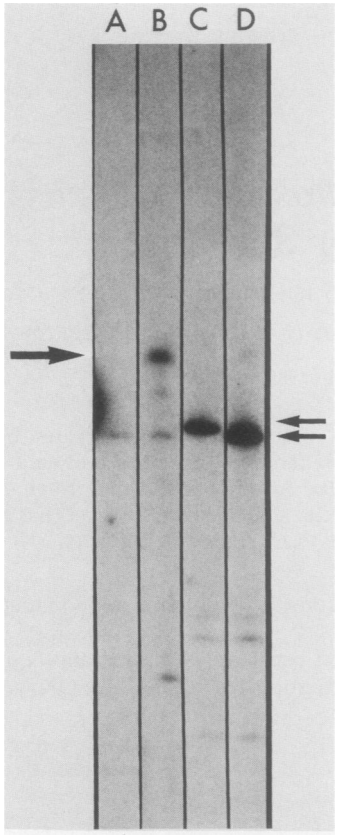

FIG. 2. Autoradiogram of hybrid-selected spliced leader transcripts and 5S rRNA transcripts from in vitro capped total trypanosome RNA. RNA was treated with periodate (lanes B and D) to remove preexisting cap structures or was untreated (lanes A and C) before capping with guanylyl transferase as above. Hybrid selection was performed with the spliced leader gene repeat (lanes A and B) or the 5S gene repeat (lanes C and D) followed by electrophoresis on a 10% acrylamide–7 M urea gel and autoradiography. The large arrow indicates the position of an RNA species 129 nucleotides in length; the upper small arrow indicates a species 118 nucleotides in length, and the lower small arrow indicates a species 117 nucleotides in length.

and used this decapped RNA as a substrate for guanylyl transferase (Fig. 1, Post). After decapping, the transcripts for the spliced leader, tubulin, and VSG were able to incorporate a much greater amount of [^{32}P]GMP, strongly suggesting that these transcripts have a preexisting cap structure resembling that of other eucaryotic mRNAs. This result is supported by the finding that the VSG mRNA can be efficiently translated in a rabbit reticulocyte lysate under salt and temperature conditions optimal for capped mRNAs (K. Presper, University of Iowa, personal communication). In contrast, the 5.8S transcript remained unlabeled after the decapping and recapping procedure. We have shown that the 5.8S RNA is likely produced from a larger precursor and has a 5′ monophosphate terminus (8, 12). Note that the capped transcripts did not bind to plasmid DNAs with no insert (Fig. 1, C1 = pUC9, C2 = pBR322) or a previously described VSG cDNA (13), MVAT 4 (Fig. 1, C3), which contained the spliced leader sequence but was not expressed in these organisms. The latter result indicates that the transcripts binding to the spliced leader are not due to the binding of degraded mRNAs which were labeled by guanylyl transferase and bear the spliced leader.

To show that specific transcripts were being labeled by this procedure, we hybrid selected the labeled species and electrophoresed them on a 10% acrylamide-urea gel (Fig. 2). Only after periodate oxidation was a spliced leader transcript of 129 ± 2 nucleotides labeled by guanylyl transferase (Fig. 2, lanes A and B). This also shows that the small amount of label which bound before decapping to the spliced leader repeat (Fig. 1) is not a labeled, full-length, spliced leader transcript. In contrast, the 5S transcript is labeled both before and after periodate treatment (12). The shortening of the 5S transcript by one nucleotide (Fig. 2, lane D) indicates that the periodate oxidation and beta elimination have successfully removed the diol-containing 3′ nucleotide. The minor 118-nucleotide labeled species in Fig. 2, lanes A and B, was present in a selection by pUC9 alone (data not shown) and probably represented labeling of a protein or DNA contaminant. These results show clearly that an ~130-nucleotide, spliced-leader-containing RNA species in total trypanosome RNA is capped.

To establish that the spliced leader transcript detected by guanylyl transferase labeling corresponded to a transcription product of the spliced leader repeat, a spliced leader transcript was sought directly. Trypanosome nuclei were prepared and transcription in vitro was carried out by modifications of the procedure described by Bitter and Roeder (1). Briefly, trypanosomes were lysed in 20 mM Tris chloride (pH 7.5)–25 mM KCl–3 mM $MgCl_2$–0.32 M sucrose–0.5% Nonidet P-40 by Dounce homogenization, washed, and suspended in a storage buffer as described previously (1). Transcription reactions of approximately 5×10^8 nuclei were carried out with the assay mixture previously described (1) for 10 min at 20°C. Incorporation was stopped by adding 10 volumes of 60% guanidinium thiocyanate–0.5% *N*-lauroyl sarcosine–50 mM lithium citrate (pH 7.0), and the RNA was isolated by centrifugation through 5.7 M cesium chloride.

The RNA thus isolated was hybrid selected with a plasmid containing the spliced leader gene repeat as previously

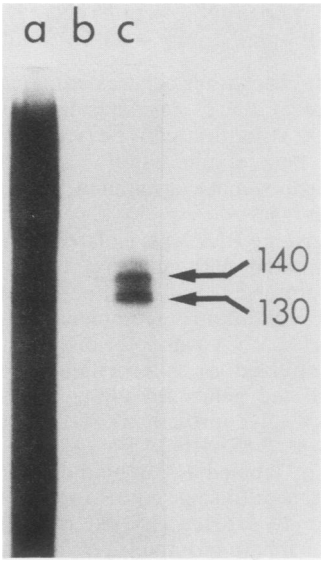

FIG. 3. Autoradiogram of a hybrid selection of transcripts encoded by the spliced leader gene repeat (a) Total unselected RNA labeled in trypanosome nuclei. (b) RNA selected by pUC9 plasmid. (c) RNA selected by spliced leader gene repeat cloned into pUC9. Arrows show approximate sizes (in nucleotides) of the two major spliced leader transcripts.

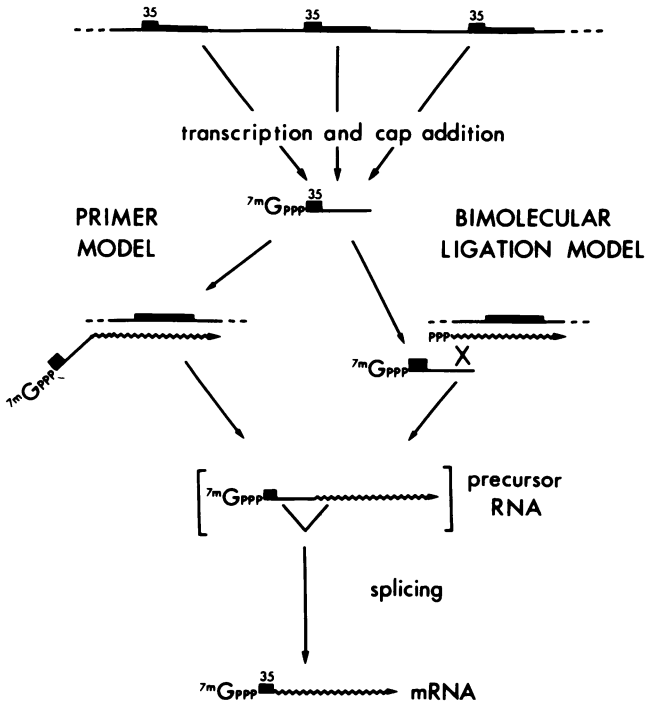

FIG. 4. Models for the spliced leader acting as a cap donor structure, based on previous models for the addition of the spliced leader sequence to protein-coding transcripts either during transcription (primer model) or as a posttranscriptional modification (biomolecular splicing model) (4, 11, 16). At the top, thick lines indicate the regions encoding the spliced leader transcripts, and doubly thick lines marked 35 are the sequence of the conserved 35-mer. Wavy lines indicate RNA transcripts which are encoded at the genomic loci for various protein-coding genes. The precursor RNA is enclosed in brackets, since its existence has not been demonstrated experimentally.

described (14). Bound RNA was eluted and electrophoresed on 10% acrylamide–7 M urea gels followed by autoradiography (Fig. 3). A complex group of spliced leader species is shown in Fig. 3, lane c; this group was synthesized with two predominant forms of ~130 and ~140 nucleotides in length (based on migration of standard-length molecules not shown). These were not selected by plasmid vector alone (lane b). No larger species were detected, suggesting that these are the primary transcripts derived from the spliced leader gene repeat. The size of these transcripts corresponded well to previous estimates of the size of the spliced leader transcript and to the finding of Kooter and Borst that a region of ~140 nucleotides of the spliced leader gene repeat is highly transcribed (4, 6, 11, 16). The significance of two major species synthesized in isolated nuclei is not clear; the smaller of the two corresponded to the size of the spliced leader transcript detected by guanylyl transferase labeling. The fact that the ~140-nucleotide species was not labeled with guanylyl transferase either before or after decapping (Fig. 2) suggests either that this species does not accumulate in steady-state RNA or that it has a 5′ monophosphate, an unusual cap structure, or a blocking molecule not sensitive to periodate oxidation. Experiments are in progress to distinguish among these possibilities. The presence of numerous minor species may be due to minor sequence variations in spliced leader RNAs produced from different mini-exon repeats within the tandem arrays, as has been previously shown for 5S rRNAs (21). Supporting this suggestion are significant differences in the sequences of the spliced leader RNA coding regions in different repeats isolated from *T. brucei* (4, 5, 7; unpublished data).

In summary, the spliced leader is a small capped RNA that bears no translated sequence (6). In this respect it is similiar to human and *Xenopus* small nuclear RNAs which are also transcribed from tandem arrays and are capped (3, 15, 22). Most previous work on the spliced leader sequence has focused on its potential role in transcription (4, 6, 11, 16, 19). Our finding that the spliced leader transcript is a capped species suggests that it also functions as a cap donor structure. Two possible models of how the spliced leader serves as a cap donor structure are shown in Fig. 4. In the primer model, an additional function of the spliced leader is to act as a primer for transcription of the mRNAs on whose 5′ end it is found. In the bimolecular ligation model, the spliced leader is added posttranscriptionally, and no predictions are made about other functions of the spliced leader sequence. A precedent supporting the notion of a small RNA acting as a cap donor structure is found with influenza virus. In this virus, the cap and a short stretch of 10 to 15 nucleotides from host cell mRNAs are removed and used as capped primers for synthesis of viral mRNAs (2). The total conservation of the 35-nucleotide sequence indicates, however, that there may be other functions served by this novel spliced leader.

This work was supported by grant DAMD 1782C2228 from the U.S. Army, Public Health Service grants AI 18954, AM 25295, and AM 07018 from the National Institutes of Health, and the Burroughs-Wellcome Foundation. D.M.D. was supported by a Medical Scientist Training Program grant from the National Institutes of Health to the University of Iowa.

LITERATURE CITED

1. **Bitter, G., and R. G. Roeder.** 1978. Transcription of viral genes by RNA polymerase II in nuclei isolated from adenovirus 2 transformed cells. Biochemistry **17**:2198–2205.
2. **Braam, J., I. Ulmansen, and R. M. Krug.** 1983. Molecular model of a eukaryotic transcription complex: functions and movements of influenza P proteins during capped RNA-primed transcription. Cell **34**:609–618.
3. **Busch, H., R. Reddy, L. Rothblum, and Y. C. Choi.** 1982. SnRNAs, SnRNPs, and RNA processing. Annu. Rev. Biochem. **51**:617–654.
4. **Campbell, D. A., D. A. Thorton, and J. C. Boothroyd.** 1984. Apparent discontinuous transcription of *Trypanosoma brucei* variant surface antigen genes. Nature (London) **311**:350–355.
5. **DeLange, T., T. Berkvans, H. G. J. Veerman, A. C. C. Frasch, J. D. Barry, and P. Borst.** 1984. Comparison of the genes coding for the common 5′ terminal sequence of messenger RNAs in three trypanosome species. Nucleic Acids Res. **12**:4431–4443.
6. **DeLange, T., A. Y. C., Liu, L. H. T. Van der Ploeg, P. Borst, M. C. Trompe, and J. H. Van Boom.** 1983. Tandem repetition of the 5′ mini-exon of varient surface glycoprotein genes: a multiple promoter for VSG transcription. Cell **34**:891–900.
7. **Dorfman, D. M., and J. E. Donelson.** 1984. Characterization of the 1.35 kilobase DNA repeat unit containing the conserved 35 nucleotides at the 5′-termini of variable surface glycoprotein mRNAs in *Trypanosoma brucei*. Nucleic Acids Res. **12**:4907–4920.
8. **Dorfman, D. M., M. Lenardo, L. V. Reddy, L. H. T. Van der Ploeg, and J. E. Donelson.** 1985. The 5.8S ribosomal RNA gene of *Trypanosoma brucei*: structural and transcriptional studies. Nucleic Acids Res. **13**:3533–3549.
9. **Fraenkel-Conrat, H., and A. Steinschneider.** 1968. Stepwise degradation of RNA: periodate oxidation followed by aniline cleavage. Methods Enzymol. **12**:243–246.
10. **Kimmel, B. E., S. Samson, R. Hirschberg, and L. Yarbrough.** 1985. Tubulin genes of the African trypanosome *Trypanosoma*

brucei rhodesiense: nucleotide sequence of a 3.7 kilobase fragment containing genes for alpha and beta tubulins. Gene **35**:237–248.

11. **Kooter, J., T. DeLange, and P. Borst.** 1984. Discontinuous synthesis of mRNA in trypanosomes. EMBO J. **3**:2387–2392.
12. **Lenardo, M. J., D. M. Dorfman, L. V. Reddy, and J. E. Donelson.** 1985. Characterization of the *Trypanosoma brucei* 5S ribosomal RNA gene and transcript: the 5S RNA is a spliced-leader-independent species. Gene **35**:131–141.
13. **Lenardo, M. J., A. C. Rice-Ficht, G. Kelly, K. Esser, and J. E. Donelson.** 1984. Characterization of the genes specifying two metacyclic variable antigen types in *Trypanosoma brucei rhodesiense*. Proc. Natl. Acad. Sci. U.S.A. **81**:6642–6646.
14. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual, p. 329–342. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **Mattaj, I. W., and R. Zeller.** 1983. *Xenopus laevis* U2 snRNA genes: tandemly repeated transcriptions units sharing 5′ and 3′ flanking homology with other RNA polymerase II transcribed genes. EMBO J. **2**:1883–1891.
16. **Milhausen, M., R. G. Nelson, S. Sather, M. Selkirk, and N. Agabian.** 1984. Identification of a small RNA containing the trypanosome spliced leader donor of the shared 5′ sequences of trypanosome mRNAs. Cell **38**:721–729.
17. **Moss, B.** 1981. 5′ end labeling of RNA with capping and methylating enzymes, p. 253–266. *In* J. G. Chirikjian and T. Papas (ed.), Gene amplification and analysis, vol. 2. Elsevier Science Publishing Co., New York.
18. **Murphy, W. J., S. T. Brentano, A. Rice-Ficht, D. M. Dorfman, and J. E. Donelson.** 1984. DNA rearrangements of the variable surface antigen genes of the trypanosomes. J. Protozool. **31**:65–73.
19. **Nelson, R. G., M. Parsons, P. J. Barr, K. Stuart, M. Selkirk, and N. Agabian.** 1983. Sequences homologous to the variant antigen mRNA spliced leader are located in tandem repeats and variable orphons in *Trypanosoma brucei*. Cell **34**:901–909.
20. **Schibler, U., O. Hagenbuchle, P. K. Wellauer, and A. C. Pittet.** 1983. Two promoters of different strengths control the transcription of the mouse alpha-amylase gene *Amy-1*^{a} in the parotid gland and the liver. Cell **33**:501–508.
21. **Selker, E. U., J. N. Stevens, and R. L. Metzenberg.** 1985. Heterogeneity of 5S RNA in fungal ribosomes. Science **227**: 1340–1343.
22. **Van Arsdell, S. W., and A. M. Weiner.** 1984. Human genes for U2 small nuclear RNA are tandemly repeated. Mol. Cell. Biol. **4**:492–499.